◇ Monograph Excerpt ◇

Matter Antimatter Fluctuations

search, discovery and analysis of Bs flavor oscillations

Nuno Leonardo

Complete work published as:
Analysis of B_s oscillations at CDF, MIT Thesis (2006)
Matter antimatter fluctuations, Monograph, LAP Lambert (2011)

Chapter 7

Measurement of B^0 oscillations

The flavor taggers are here applied to the $B^{+,0}$ data samples, the b-flavor information is incorporated in the fit, and the frequency of the oscillations in the B^0 system is extracted.

7.1 Fitting technique

The analysis of B flavor oscillations is performed using the unbinned maximum likelihood estimation method. The fitting framework and technique are built upon those implemented in Chapter 5 for the analysis of mass and proper decay time in the absence of flavor tagging information. The latter constitutes the ingredient which needs now to be incorporated.

The likelihood function is as in (5.4)

$$\mathcal{L} = \prod_i \sum_\alpha f_\alpha \mathcal{P}_i^\alpha \tag{7.1}$$

given by the product over the sample candidates (denoted by the index i) of the combination of the likelihood terms describing each of the sample components (denoted by the index α).

Besides the fit input quantities that have already been introduced in Section 5.1.2 – mass m, proper decay time t, and uncertainty σ_t – the novel information to be provided to the fit is the *decision* ξ and the expected *dilution* $\mathcal{D}$ of the tagging methods evaluated for the individual B meson candidates. The tagging decision constitutes the comparison of the flavor states of the B candidate at production and decay times performed by the tagging methods, while the tagging dilution is related to the correctness of the decision.

The likelihood terms describing each sample component, $\mathcal{P}^\alpha$, are given by the joint PDFs of the fit input variables,

$$\mathcal{P} \;=\; L_m \, L_t \, L_\xi \, L_\mathcal{D} \, L_{\sigma_t}. \tag{7.2}$$

The factorization given in (7.2) along with the meaning of the involved likelihood factors can be derived similarly to what was done in (5.12) using the rules of conditional probability, and relative dependences among the variables, as follows

$$\begin{aligned}\mathcal{P}(m,t,\sigma_t,\xi,\mathcal{D}) &= \mathcal{P}(m)\cdot\mathcal{P}(t,\xi|\mathcal{D},\sigma_t)\cdot\mathcal{P}(\mathcal{D})\cdot\mathcal{P}(\sigma_t) \\ &= \mathcal{P}(m)\cdot\mathcal{P}(t|\xi,\mathcal{D},\sigma_t)\cdot\mathcal{P}(\xi|\mathcal{D},\sigma_t)\cdot\mathcal{P}(\mathcal{D})\cdot\mathcal{P}(\sigma_t)\ . \end{aligned} \tag{7.3}$$

The mass is decoupled from the functional dependence of probabilities on the other observables. The description in proper decay time space is going to be determined by the candidate lifetime along with mixing, tagging and resolution effects. In general, it will be distinct for events belonging to different tag classes, as specified by the tagging decision, and will depend on the candidate dilution in addition to its proper decay time resolution.

The likelihood factors L_{σ_t} and $L_{\mathcal{D}}$ do not involve likelihood parameters and are given instead by distributions obtained directly from the data samples. For each of the involved input quantities, namely the uncertainty σ_t and the dilutions $\mathcal{D}$ for the various taggers, two distributions are constructed: one from mass-sideband candidates and the other from the mass signal region after mass-sideband subtraction. These are respectively associated to the combinatorial background and the remaining sample components. The actual distribution used for $L_{\mathcal{D}}$ is that associated to the tagging method which provides the tagging decision ξ for the candidate at hand.

The description of the mass probability distributions L_m for the various sample components is addressed in Section 5.4. The likelihood models for proper decay time and flavor tagging contain the parameters of interest to be extracted from the fit to the data samples, and are thoroughly derived below.

7.2 Flavor tagging information

Several b-flavor tagging methods are applied to the B meson candidates. These were presented in Section 6. The class of algorithms denoted opposite-side taggers, in brief OST, is applied to all samples, and the same-side tagger, SST, method is applied solely to the fully reconstructed $B^{+,0}$ candidates.

The tagging information is contained in the following quantities: decision ξ, dilution $\mathcal{D}$, and efficiency ϵ. The tagging decision takes on the discrete values $\xi \in \{+1,-1,0\}$. A positive (negative) decision indicates that the tagged b-flavor at the time the B meson was produced coincides with (is opposite to) the b-flavor at the time of decay. A null decision is assigned whenever the tagger fails to identify the b flavor at production time. The fraction of events for which a non-trivial decision ($\xi \neq 0$) is achieved defines the tagging efficiency, while the

probability that such a decision is correct is determined by the tagging dilution,

$$\mathcal{P}(\text{correct decision}) = \frac{1+\mathcal{D}}{2}\,.$$

The OSTs expected dilution is evaluated for each candidate, and provided as input data to the fit along with the SST and OST decisions.

The various OST methods are combined in an exclusive way. Specifically, the taggers are hierarchically ranked based on expected performance; whenever non-trivial decisions $\xi = \pm 1$ are achieved for a given candidate by a set of methods, only that provided by the foremost is used. This is further elucidated in Table 7.1.

opposite-side method	abbreviation	decision hierarchy
soft muon	SMT	evaluated first
soft electron	SET	if SMT failed
jet charge, class 1	JVX	if SMT and SET failed
jet charge, class 2	JJP	if SMT, SET and JVX failed
jet charge, class 3	JPT	if SMT, SET, JVX and JJP failed

Table 7.1: List of opposite-side taggers and decision hierarchy.

7.3 Likelihood formalism for flavor tagged samples

The likelihood descriptions of proper decay time and flavor tagging spaces are in general interconnected, and are more conveniently summarized by the joint PDF

$$L_{t,\xi} \;=\; \mathcal{P}(t,\xi|\mathcal{D},\sigma_t)\,. \tag{7.4}$$

As it will be shown, in the case of sample components which do not undergo flavor oscillations, the proper time PDF may be defined in a decoupled fashion from the tagging decision information. It will then coincide with the t-PDFs which were derived in Chapter 5, in the absence of flavor tagging information. The probability of observing a given flavor tagging decision will in that case be determined exclusively from quantities which characterize the flavor tagging methods when applied to the data samples – the tagging efficiency and dilution.

For those components describing neutral B meson candidates a cosine term is included for characterizing the time-dependent flavor oscillations. Correspondingly, the probability for a given tagging decision to be observed will depend on that for mixing to have occurred, and will thus have a t dependence.

We proceed by first considering the simpler case of a single tagging algorithm, before extending the derived formalism to the actual situation of several taggers.

7.3.1 Non-oscillating components

We first address those sample components for which mixing effects do not occur. This is the model used for describing charged B meson signals, as well as backgrounds such as those of combinatorial type. In this case proper time and tagging probabilities may be treated in a decoupled fashion,

$$L_{t,\xi} \;=\; \mathcal{P}(\xi|\mathcal{D}) \cdot \mathcal{P}(t|\sigma_t) = L_\xi \, L_t \,. \tag{7.5}$$

The proper decay time PDF, L_t, coincides in this case with the modeling derived in Chapter 5. The tagging likelihood term L_ξ is new and needs now to be addressed.

The probability that an event, as associated to a given sample component, is tagged is expressed by the tagging efficiency ϵ for that component, which is a parameter of the fit. The non-tagging ($\xi = 0$) probability is then given by $1 - \epsilon$. This is summarized by the factor p_ϵ defined as

$$p_\epsilon(\xi) = \mathcal{P}(\xi|\epsilon) \;=\; \epsilon|\xi| + (1-\epsilon)(1-|\xi|) = \left\{ \begin{array}{cl} 1-\epsilon & \text{for } \xi = \;\; 0\,, \\ \epsilon & \text{for } \xi = \pm 1\,. \end{array} \right. \tag{7.6}$$

In case a non-trivial tagging decision is achieved, the probability that it is correct is determined by the tagging dilution evaluated for the candidate, and is given by $\frac{1+\mathcal{D}}{2}$. Provided no oscillations take place, as we are assuming to be the case, the indication of flavor change ($\xi = -1$) must be due to mistagging, whose probability is $\frac{1-\mathcal{D}}{2}$. This can be summarized as

$$L_\xi(\xi|\mathcal{D},\epsilon) \;=\; p_\epsilon(\xi) \cdot \frac{1+\xi\mathcal{D}}{1+|\xi|} = \left\{ \begin{array}{cl} 1-\epsilon & \text{for } \xi = \;\; 0\,, \\ \epsilon \; \frac{1\pm\mathcal{D}}{2} & \text{for } \xi = \pm 1\,; \end{array} \right. \tag{7.7}$$

where we note that the necessary normalization $\sum_{\xi\in\{0,+1,-1\}} L(\xi) = 1$ is verified.

7.3.2 Oscillating signal

When mixing is present, specific proper decay time distributions characterize the samples associated to the different tagging decisions. The probability for a flavor change to be ($\xi = -1$) or not be ($\xi = +1$) verified, at true proper time t, may be expressed as

$$\mathcal{P}(\xi|t,w) = \frac{1+\xi\cos wt}{2}\,, \tag{7.8}$$

where $w = \Delta m$ is the oscillation frequency of the system. The effect of mistagging is readily evaluated as

$$\mathcal{P}(\xi|t,\mathcal{D},w) \;=\; \frac{1+\xi\cos wt}{2}\frac{1+\mathcal{D}}{2} + \frac{1-\xi\cos wt}{2}\frac{1-\mathcal{D}}{2} = \frac{1+\xi\mathcal{D}\cos wt}{2}\,, \qquad (7.9)$$

and additionally, using (7.6), so is that of tagging inefficiency

$$\mathcal{P}(\xi|t,\mathcal{D},\epsilon,w) \;=\; p_\epsilon(\xi)\cdot\frac{1+\xi\mathcal{D}\cos wt}{1+|\xi|} = \begin{cases} 1-\epsilon & \text{for } \xi = 0\,, \\ \epsilon\,\frac{1\pm\mathcal{D}\cos wt}{2} & \text{for } \xi = \pm 1\,. \end{cases} \qquad (7.10)$$

The joint PDF for true proper decay time and tagging decision is given by $\mathcal{P}(t,\xi|\mathcal{D},\epsilon,w) = \mathcal{P}(\xi|t,\mathcal{D},\epsilon,w)\;\mathcal{P}(t)$, where the proper time probability decoupled from tagging information (that is integrated over the tagging decisions) has the lifetime exponential decay form of (5.13),

$$\begin{aligned}\mathcal{P}(t,\xi|\mathcal{D},\epsilon,w,\tau) &= p_\epsilon(\xi)\cdot\frac{1+\xi\mathcal{D}\cos wt}{1+|\xi|}\frac{1}{\tau}e^{-\frac{t}{\tau}}\theta(t) \\ &= p_\epsilon(\xi)\cdot\frac{E(t;\tau)+\xi\mathcal{D}\,C(t;\tau,w)}{1+|\xi|}\,, \end{aligned} \qquad (7.11)$$

where we have for convenience introduced the functions E and C, whose definitions are implicitly given by the second equality (as in (5.13)).

For partially reconstructed modes, the pseudo proper decay time needs to be corrected by a kinematical factor, κ, defined in (5.8). This correction is achieved through the smearing of (7.11) with the κ-factor distribution $\mathcal{F}(\kappa)$ (Section 5.2.3). Specifically, the functions E and C become modified as follows

$$\begin{aligned} E'(t;\tau)+\xi\mathcal{D}\;C'(t;\tau,w) &= [E(t;\tau)+\xi\mathcal{D}\;C(t;\tau,w)]\otimes_\kappa\mathcal{F}(\kappa) \\ &= \int(1+\xi\mathcal{D}\cos(w\kappa t))\,\frac{\kappa}{\tau}e^{-\frac{\kappa t}{\tau}}\theta(\kappa t)\cdot\mathcal{F}(\kappa)\;d\kappa\,, \end{aligned} \qquad (7.12)$$

where the κ-smearing operator $\otimes_\kappa$, originally defined in (5.21), has here been extended.

The effects of detector resolution, trigger and selection requirements are addressed as well in Section 5.2. With the description of such effects included, the joint PDF can accordingly be expressed as

$$L_{t,\xi}(t,\xi|\mathcal{D},\epsilon,w,\tau,\sigma_t) = p_\epsilon(\xi)\frac{1}{\mathcal{N}}\frac{E(t;\tau)+\xi\mathcal{D}\;C(t;\tau,w)}{1+|\xi|}\otimes_\kappa\mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\,, \qquad (7.13)$$

where $G(t;\sigma_t)$ is the proper decay time resolution function (5.14), and $\mathcal{E}(t)$is the t-efficiency function defined in Section 5.3.1. The normalization constant $\mathcal{N}$ is given in (5.20) and (5.23). The PDF normalization condition $\sum_\xi\int L(t,\xi)dt = 1$ is verified.

To conclude we mention that (7.13) may be cast in a form which explicitly verifies the factorization expressed in (7.2),

$$\begin{aligned} L_{t,\xi}(t,\xi) &= L_t(t|\xi)\cdot L_\xi(\xi)\ , \\ L_t(t|\xi) &= \frac{1}{N_\xi}\,\frac{E(t;\tau)+\xi\mathcal{D}\;C(t;\tau,w)}{1+|\xi|}\otimes_\kappa \mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\ , \\ L_\xi(\xi) &= f_\xi\ . \end{aligned} \tag{7.14}$$

The factors N_ξ are found by evaluating the normalization condition $\int L_t(t|\xi)dt = 1$ for each decision. The factors f_ξ, to be determined, must similarly ensure the corresponding normalization condition $\sum_{\xi\in\{0,-1,+1\}} L_\xi(\xi) = 1$. Furthermore, they correspond to the relative fractions of the various tagging decisions. The requirements $f_0 = 1-\epsilon$, $f_+ + f_- = \epsilon$, and $f_+/f_- = N_+/N_-$ readily imply

$$f_\xi \;=\; [\epsilon|\xi| + (1-\epsilon)(1-|\xi|)]\cdot\frac{N_\xi}{\mathcal{N}}\ , \tag{7.15}$$

and (7.13) is recovered. We in particular emphasize that the evaluation of the factors N_ξ which would involve the computation of integrals of cosine terms are not necessary. Solely the normalization factor $\mathcal{N}$which was originally addressed in the context of the lifetime analysis is required.

7.3.3 Multiply-tagged candidates

We address now the case in which several decisions, provided by different tagging methods, are available for a single event. In case these decisions are correlated – as between those provided by different OSTs – one of the algorithms is selected, while for uncorrelated taggers (as between OST and SST) the information provided by the algorithms is fully combined.

For candidates concurrently tagged by several OST algorithms, correspondingly characterized by multiple decisions $\{\xi_i\}$ and dilutions $\{\mathcal{D}_i\}$, only the information provided by the selected one among these, as it was described in Section 7.2, is employed in the fit. Each algorithm is in this way effectively employed to tag the flavor of a fraction of candidates. These fractions are denoted by the parameters $\{\epsilon_i\}$. An event being not tagged would imply a trivial decision $\xi_i = 0$ among all algorithms, which occurs for a fraction $(1-\sum_i \epsilon_i)$ of the candidates. Let the selected algorithm be identified by the index j. The factor of (7.6) becomes then

$$p_{\{\epsilon_i\}}(\xi_j) = \epsilon_j|\xi_j| + (1-\sum_i \epsilon_i)(1-|\xi_j|) = \begin{cases} 1-\sum_i \epsilon_i & \text{for}\quad \forall_i\,\xi_i = 0\ , \\ \epsilon_j & \text{for}\quad \xi_j = \pm 1\ . \end{cases} \tag{7.16}$$

The PDF (7.13) becomes accordingly generalized to

$$\begin{aligned} &L_{t,\xi}(t,\xi_j|\mathcal{D}_j,\{\epsilon_i\},w,\tau,\sigma_t) \\ &\quad = p_{\{\epsilon_i\}}(\xi_j)\cdot\frac{1}{\mathcal{N}}\,\frac{E(t;\tau)+\xi_j\mathcal{D}_j\;C(t;\tau,w)}{1+|\xi_j|}\otimes_\kappa\mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\;. \end{aligned} \tag{7.17}$$

When information provided by both the SST algorithm and a selected OST method is to be used in the fit, the combination of such information needs to be implemented in the likelihood model. The issue of combining independent taggers is explored in Section 6.3. The parameters ξ', $\mathcal{D}'$, and ϵ' are used to characterize the SST, while unprimed symbols continue to be used for the OSTs. The efficiency factors are given simply by the multiplication

$$p_{\{\epsilon_i\}}(\xi_j)\cdot p_{\epsilon'}(\xi') \;=\; \begin{cases} (1-\sum_i\epsilon_i)\cdot(1-\epsilon') & \text{for}\;\; \forall_i\,\xi_i = \;\;0,\;\; \xi' = \;\;0\;, \\ (1-\sum_i\epsilon_i)\cdot\epsilon' & \text{for}\;\; \forall_i\,\xi_i = \;\;0,\;\; \xi' = \pm1\;, \\ \epsilon_j\cdot(1-\epsilon') & \text{for}\;\; \xi_j = \pm1,\;\; \xi' = \;\;0\;, \\ \epsilon_j\cdot\epsilon' & \text{for}\;\; \xi_j = \pm1,\;\; \xi' = \pm1\;. \end{cases} \tag{7.18}$$

The (t,ξ)-PDF takes in this case the following form

$$\begin{aligned} L_{t,\xi}(t,\xi_j,\xi'|\mathcal{D}_j,\{\epsilon_i\},\mathcal{D}',\epsilon',w,\tau,\sigma_t) \;&=\; p_{\{\epsilon_i\}}(\xi_j)\cdot p_{\epsilon'}(\xi') \\ \times\frac{1}{\mathcal{N}}\frac{(1+\xi_j\xi'\mathcal{D}_j\mathcal{D}')\,E(t;\tau)+(\xi_j\mathcal{D}_j+\xi'\mathcal{D}')\,C(t;\tau,w)}{(1+|\xi_j|)(1+|\xi'|)}&\otimes_\kappa\mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\;. \end{aligned} \tag{7.19}$$

The joint proper decay time and tagging PDF of (7.19) is written in a most general form for the signal likelihood components. It straightforwardly reduces to the cases of lesser taggers derived before. It is suitable directly for the partially-reconstructed, t-biased samples. Its form for describing the fully reconstructed samples is recovered by imposing $\mathcal{F}(\kappa)=\delta(\kappa-1)$, and, for the t-unbiased samples, $\mathcal{E}(t)=1$. Incidentally, it may also be seen to hold for both neutral and charged B meson decay modes, with the parameter identification $w=\Delta m$ and by fixing $w=0$, respectively.

For components where mixing does not participate, or it is not to be explicitly parameterized, the descriptions of tagging and proper time is performed in a decoupled fashion, (7.5). The t-PDF coincides in those cases with that given in Chapter 5 for the sample component at hand. In adddition, so does the description of each of the independent flavor taggers. The joint (t,ξ)-PDF may be expressed as follows

$$\begin{aligned} L_{t,\xi}(t,\xi_j,\xi'|\mathcal{D}_j,\{\epsilon_i\},\mathcal{D}',\epsilon') \;&=\; L_\xi(\xi_j|\mathcal{D}_j,\{\epsilon_i\})\cdot L_\xi(\xi'|\mathcal{D}',\epsilon')\cdot L_t(t) \\ &=\; p_{\{\epsilon_i\}}(\xi_j)p_{\epsilon'}(\xi')\cdot\frac{1+\xi_j\mathcal{D}_j}{1+|\xi_j|}\frac{1+\xi'\mathcal{D}'}{1+|\xi'|}\cdot L_t(t)\;. \end{aligned} \tag{7.20}$$

7.4 Mixing and dilution calibration

In the fits of the B^0 and B^+ data samples described in the current chapter, we are interested in measuring the oscillation frequency Δm_d along with the taggers performance. The former appears as argument of the cosine function describing flavor oscillations as part of the signal PDF for neutral modes, while the taggers dilutions are determined, simultaneously, from both charged and neutral signal components.

The dilution of the OSTs is predicted, and assigned to each candidate as fit input, based on the parameterizations from Chapter 6. These parameterizations are achieved in high statistics samples with high B signal purity, and provide an optimal tagging power by distinguishing among candidates with poorer and higher dilutions. Accordingly, in place of extracting from the fit an average dilution for each OST, we use the referred per-event predicted dilution values, and allow instead for overall calibration factors. A single scale factor S_i is employed for each individual opposite side tagger, as a fit parameter multiplying the corresponding predicted dilution in the likelihood model.

The introduction of such dilution scale factors allows for the quantification of differences between the predicted and actual dilutions of the OST algorithms applied to our signal samples [61]. If the dilution parameterizations are adequate and directly applicable to the signal samples to be fitted, the scale factors are then expected to be consistent with unit.

Furthermore, the behavior of the OSTs is expected to be identical when applied to samples of different B meson species. The calibration of the OSTs dilution through determination of the dilution scale factors is thus among the most interesting results to be presently obtained. These calibrations obtained based on the fully and partially reconstructed samples will be transferred directly to the kinematically and topologically similar B_s samples, where flavor oscillations will be studied in Chapter 8. Then, the dilution scale factors are provided as necessary input information to the likelihood model.

The joint proper time and tagging decision PDFs for the signal components derived in the previous section are modified slightly to include the dilution scale factor parameters $\{S_i\}$. For the partially reconstructed B^+ and B^0 signals in the DlX samples, following (7.17), it takes the form

$$
\begin{aligned}
L_{t,\xi}(t,\xi_j|\,S_j\,,\mathcal{D}_j,\{\epsilon_i\},\Delta m_d,\tau_u,\tau_d,\sigma_t) \;&=\; p_{\{\epsilon_i\}}(\xi_j)\,\frac{1}{1+|\xi_j|} \\
&\times\,\{\; f_u\cdot\frac{1}{\mathcal{N}_u}\,[1+\xi_j\,S_j\,\mathcal{D}_j]\,E(t;\tau_u)\otimes_\kappa\mathcal{F}_u(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}_u(t) \\
&+f_d\cdot\frac{1}{\mathcal{N}_d}\,[E(t;\tau_d)+\xi_j\,S_j\,\mathcal{D}_j\;C(t;\tau_d,\Delta m_d)]\otimes_\kappa\mathcal{F}_d(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}_d(t)\;\}\;,
\end{aligned}
\tag{7.21}
$$

where the $u(d)$-indexed quantities refer to the $B^+(B^0)$ components. In this expression the

expected common OST behavior for different B species is manifest: the same dilution and scale factor parameter are used in the B^+ and B^0 PDF terms.

For the fully reconstructed samples, in addition to the OST algorithms the SST method is also used. For the SST, no dilution information is provided as input to the fit. Instead, an average dilution $\mathcal{D}'$ is included in the likelihood model as a fit parameter. The joint tagging and proper time PDF for the fully reconstructed signal components is given by (7.19) with the inclusion of the OST dilution scale factors,

$$\begin{aligned} L_{t,\xi}(t,\xi_j,\xi'|S_j,\mathcal{D}_j,\{\epsilon_i\},\mathcal{D}',\epsilon',w,\tau,\sigma_t) \;&=\; p_{\{\epsilon_i\}}(\xi_j)\frac{1}{1+|\xi_j|}\cdot p_{\epsilon'}(\xi')\frac{1}{1+|\xi'|} \\ \times\frac{1}{\mathcal{N}}\,\{\,(1+\xi_j\xi' S_j\mathcal{D}_j\mathcal{D}')\,E(t;\tau)+(\xi_j S_j\mathcal{D}_j+\xi'\mathcal{D}')\,C(t;\tau,w)\,\}\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\;. \end{aligned} \tag{7.22}$$

We note that, contrarily to the OST case, the SST properties are expected to be distinct for different B species; *e.g.* the value of the dilution parameter $\mathcal{D}'$ is anticipated to substantially differ when fitting B^+ and B^0 meson samples.

Sample components describing B decays other than signal are in general assigned the same tagging parameters as signal. In particular, the Cabibbo-suppressed decays share the same tagging and proper decay time description as the signal.

For non-physics background components the predicted OST dilution is meaningless, and it is thus not used to describe the tagging or proper time PDFs. In order to account for possible background flavor asymmetries, the tagging decision(s) are used in the background likelihood modeling together with dilution-like parameters, $\mathcal{D}^B$. These will be referred to as background dilution parameters, and one such parameter is associated to each of the (OST and SST) tagging methods, for each sample. No time-dependent asymmetries are assumed, nor expected, and the model of (7.20) is employed, for both combinatorial and fakes backgrounds.

The proper decay time models L_t for the sample components other than nominal signals where mixing is present (*i.e.* B^0 modes) coincide with those found in Chapter 5. The likelihood models in mass space L_m for all the individual components of the various data samples are also those described therein. The remaining likelihood factors, L_{σ_t} and $L_{\mathcal{D}}$, do not explicitly involve fit parameters and are given by distributions obtained as described in Section 7.1. Some further considerations related to the treatment of flavor tagging information that are specific to individual samples are provided below.

$B^{+,0} \to J/\psi K$

The Cabibbo-suppressed $B^+ \to J/\psi\pi^+$ decays share the same proper time and tagging model with the nominal signal.

The K^*-swap component deserves special treatment, because the decay flavor for these candidates is in fact opposite to that indicated by the reconstruction. The B flavor at decay time is inferred from the charges of the $K^{*0} \to K^+\pi^-$ daughter particles, and the misassignement of kaon and pion hypotheses leads to an incorrect inference of the actual B meson decay flavor. For tagged events, therefore, the taggers decisions are flipped for the swap component, which otherwise shares the same proper decay time and tagging descriptions as those used for signal.

$B^{+,0} \to D^{(*)}\pi(\pi\pi)$

The proper time and tagging joint PDF for the nominal signal component is given by (7.22), which is shared by the associated Cabibbo-suppressed $B \to D^{(*)}K(\pi\pi)$ decay component.

In the sample of $B^0 \to D^-\pi^+$ candidates, the B_s and Λ_b^0 background components are described through models of the form of (7.20), where the proper decay time is modeled in identically the same fashion as was used in the absence of flavor tagging information addressed in Chapter 5. For the misreconstructed B_s decays, no global flavor asymmetry is considered, considering the rapid anticipated oscillations. For the Λ_b^0 component the tagging parameters are common to those used for the signal, except for the SST dilution which has not been measured and to which a value of 16% is assigned. A systematic variation of this value will be considered as part of the uncertainty.

For the $D^*\pi(\pi\pi)$ samples, the dilution template distributions $L_{\mathcal{D}}$ are derived commonly for all four modes as these are affected by limited statistics.

$B^{+,0} \to DlX$

The signal components of the semileptonic samples are described by the model derived in (7.21). The remaining components are modeled in the proper decay time space as in Chapter 5, while the tagging parameters are contained in the likelihood factor L_ξ given by the general prototype of (7.7).

For the physics backgrounds, the OST dilution is given for each candidate by its predicted value, while the tagging parameters $\{S_i\}$ and $\{\epsilon_i\}$ coincide with those used for describing the signal components.

The description of the combinatorial and fakes background components does not make use of the predicted OST dilution input values, and are each described by an independent

set of flavor asymmetry $\{\mathcal{D}_i\}$ and efficiency $\{\epsilon_i\}$ parameters.

7.5 Fitting procedure and results

The model parameters describing the mass and proper time space for our samples have been obtained in the fits to the data performed in Chapter 5 in the absence of flavor tagging information. At present, the aim is to determine those parameters introduced in the likelihood models that are associated to tagging and mixing. These correspond to:

1. the taggers efficiencies, $\{\epsilon_i\}$ and ϵ', for backgrounds and signal components,

2. the flavor asymmetries of non-physics backgrounds for each tagger,

3. the OST dilution scale factors $\{S_i\}$ and (for the fully reconstructed modes) the SST dilutions, for signal and physics backgrounds components,

4. the mixing frequency Δm_d, for the B^0 modes.

A characterization of the non-physics background components is achieved by fitting the individual data samples separately. The taggers efficiencies and flavor asymmetries for the combinatorial background are determined from fits to candidates in the mass-sideband region. The efficiencies for the fakes background component in the semileptonic modes are constrained to those of the signal, and its flavor asymmetry parameters are determined from the full fit to the individual samples.

Finally, combined fits are performed for the fully and the partially reconstructed B samples separately to determine the tagging parameters which describe both signal and physics backgrounds, as well as the B^0 mixing frequency. In such combined fits the likelihood maximization is performed on the product of the likelihoods of the individual samples according to (5.2), which are functions of those common parameters.

In the fit to the combination of partially reconstructed B samples, a simultaneous determination of the taggers signal efficiencies ϵ_i, the OSTs dilution scale factors S_i, and the mixing parameter Δm_d are achieved. In the fit of the fully reconstructed modes the SST dilutions for B^+ and B^0 are in addition determined.

The fitted values of these parameters are presented in Table 7.2. For the main parameters of interest – OST dilution scale factors, SST dilutions, and oscillation frequency – the systematic uncertainties are also shown, which are evaluated in Section 7.6.

Fit projections are shown for the high statistics, semileptonic samples in Figure 7.1. These are obtained for each individual OST method, as the tagged flavor, time-dependent

parameter	hadronic fit	semileptonic fit
ϵ_{SMT}	0.0458 ± 0.0013	0.04808 ± 0.00044
ϵ_{SET}	0.0261 ± 0.0009	0.02959 ± 0.00036
ϵ_{JVX}	0.0821 ± 0.0016	0.08208 ± 0.00057
ϵ_{JJP}	0.2837 ± 0.0027	0.27578 ± 0.00093
ϵ_{JPT}	0.5146 ± 0.0030	0.51626 ± 0.00104
ϵ'_{SST}	0.6339 ± 0.0028	–
S_{SMT}	$0.95 \pm 0.08 \pm 0.02$	$0.936 \pm 0.038 \pm 0.013$
S_{SET}	$1.06 \pm 0.11 \pm 0.05$	$1.072 \pm 0.045 \pm 0.017$
S_{JVX}	$0.88 \pm 0.13 \pm 0.03$	$0.917 \pm 0.058 \pm 0.013$
S_{JJP}	$0.98 \pm 0.12 \pm 0.03$	$1.001 \pm 0.060 \pm 0.028$
S_{JPT}	$1.20 \pm 0.22 \pm 0.09$	$0.838 \pm 0.106 \pm 0.043$
$\mathcal{D}'_{SST}(B^+)$	$0.209 \pm 0.011 \pm 0.001$	–
$\mathcal{D}'_{SST}(B^0)$	$0.130 \pm 0.018 \pm 0.003$	–
Δm_d [ps^{-1}]	$0.536 \pm 0.028 \pm 0.006$	$0.511 \pm 0.020 \pm 0.014$

Table 7.2: Tagging and mixing results from the combined fits in the fully and the partially reconstructed B^+ and B^0 samples; the first uncertainty is statistical and the second systematic.

asymmetry for signal. Denoting by $N_\xi(t)$ the pseudo proper decay time distribution of events for which the tagging decision is ξ, the shown asymmetry is constructed as

$$\frac{N_+(t) - N_-(t)}{N_+(t) + N_-(t)} \,. \tag{7.23}$$

For the likelihood projections themselves, $N_\xi(t)$ are based on (7.21). The figures also show these asymmetries obtained for the B^+ and B^0 terms separately, characterized respectively by a flat and an oscillating pattern. For the points representing the data in those figures, the proper time distributions $N_\xi(t)$ are obtained from the sub-samples of candidates characterized by the different tagging decisions, after subtracting background contributions using the associated fit models, and evaluating the corresponding asymmetries in the introduced t-bins.

It is also interesting to evaluate the overall tagging effectiveness, $\epsilon\mathcal{D}^2$. For the SST this quantity is readily computed from the combined fit results, while for the OSTs it is estimated by

$$\epsilon S^2 \left\langle \mathcal{D}^2 \right\rangle$$

where the average over per-event predicted dilutions is evaluated, and further mass-sideband subtracted. The results are found in Table 7.3.

tagger	$\epsilon\mathcal{D}^2$ [%]	
	hadronic	semileptonic
SMT	$0.559 \pm 0.094 \pm 0.027$	$0.551 \pm 0.048 \pm 0.020$
SET	$0.264 \pm 0.054 \pm 0.022$	$0.308 \pm 0.030 \pm 0.008$
JVX	$0.230 \pm 0.068 \pm 0.017$	$0.247 \pm 0.033 \pm 0.010$
JJP	$0.347 \pm 0.084 \pm 0.020$	$0.366 \pm 0.045 \pm 0.021$
JPT	$0.152 \pm 0.055 \pm 0.024$	$0.076 \pm 0.019 \pm 0.009$
total OST	$1.553 \pm 0.163 \pm 0.050$	$1.550 \pm 0.083 \pm 0.029$
SST, B^0	$1.074 \pm 0.302 \pm 0.050$	—
SST, B^+	$2.773 \pm 0.296 \pm 0.027$	—

Table 7.3: Tagging effectiveness $\epsilon\mathcal{D}^2$ for each tagging method; the first uncertainty is statistical and the second systematic.

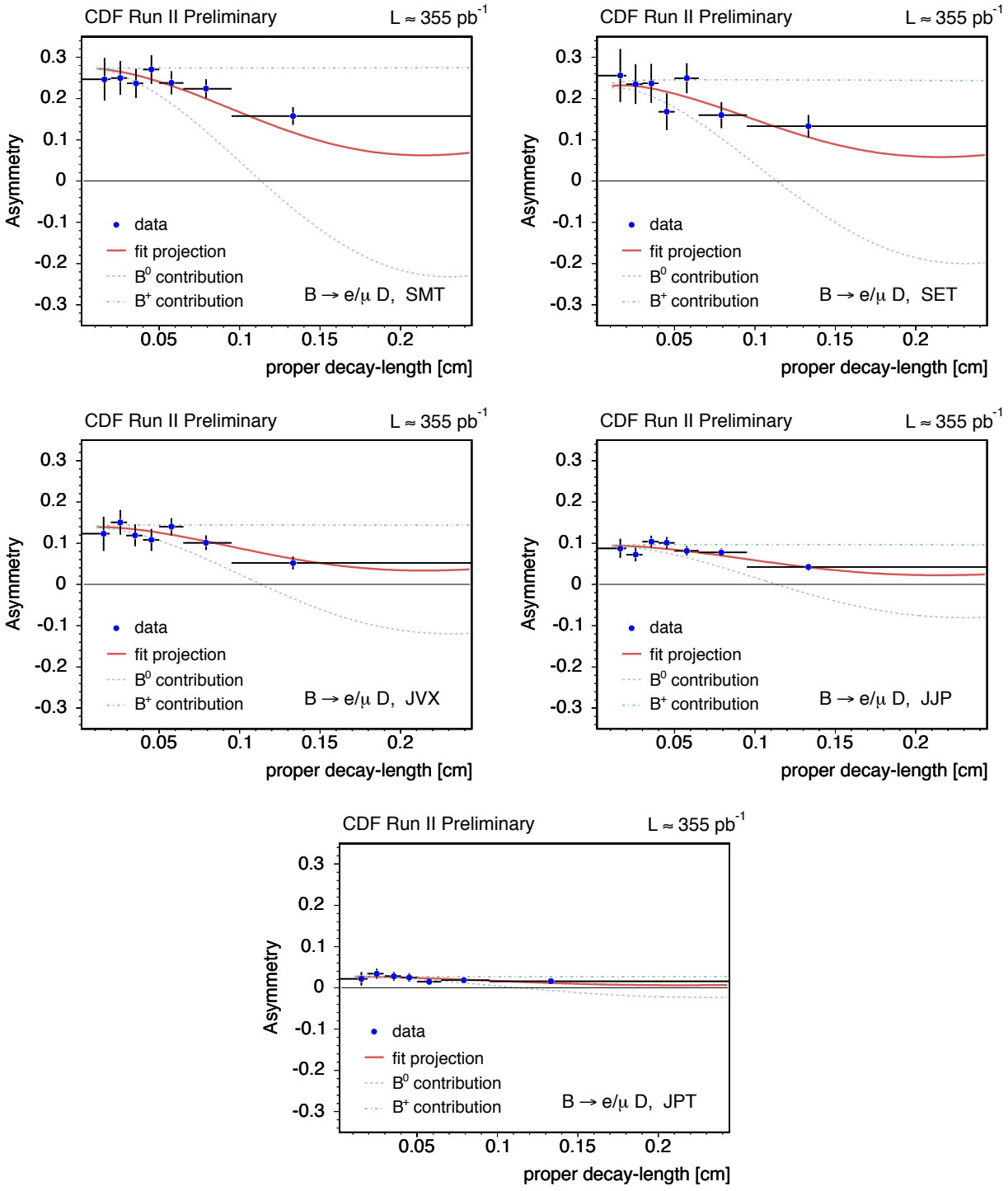

Figure 7.1: Asymmetry projections of the combined semileptonic fit for individual OST methods.

7.6 Systematic uncertainties

Systematic uncertainties are evaluated for the oscillation frequency Δm_d, as well as for the OST dilutions scale factors and the SST dilutions, by repeating the fits under modified conditions dictated by the considered sources enumerated below.

Combinatorial background m and t template parameters: The description of several background components is pre-determined and fixed in the final combined fits. Modifications of these parameters are propagated into systematic uncertainties for our primary parameters of interest. For the $D\pi$ modes, mass parameters for the combinatorial background component are pre-determined from a wide-range mass fit. They have uncertainties which are translated into systematic uncertainties by varying their values within $\pm 1\sigma$ and re-running the combined fit. For the semileptonic modes, such parameter variations are instead achieved by performing the fits in mass-sideband regions extended by ± 50 MeV/c^2. For the fully reconstructed samples, additionally, those parameters which are fixed to the values found in the fits of the individual modes are smeared according to the corresponding uncertainties. Specifically, such uncertainties are used as the widths of Gaussian-distributed corrections, which are added for each such parameter to its nominally fixed value. The combined fit is repeated many (about 150) times, and the resulting distributions of the parameters of interest are crudely Gaussian, the widths of these distributions being taken as estimates of systematic uncertainty.

Physics background levels: Various background component contributions are evaluated from Monte Carlo simulation and further specific branching ratios information. The fractions of Cabibbo-suppressed components in $D\pi\pi\pi$ samples are reset to 0-12% of the associated signal component contribution. For the $B^0 \to J/\psi K^{*0}$ mode, the fraction of events with swapped $K\pi$ mass assignment in K^{*0} candidates relative to that of correctly reconstructed signal candidates is re-estimated from an independent Monte Carlo sample. The modified relative fraction value thus obtained is 19%. In the $B^0 \to D^-\pi^+$ sample, the amounts of Λ_b^0 and B_s background decays are varied by 30%. For the semileptonic samples various branching ratios are involved in the determination of the physics background fraction which are poorly measured. Fractional variations of $\pm 25\%$ relative uncertainty are considered.

Semileptonic signal composition: Each of the three reconstructed Dl signal final states involve multiple decay chains as discussed in Section 4.3. The sample composition parameters for the B^+ and the B^0 signals are varied within the associated uncertainties. For the parameter P_V, for which no direct measurement yet exists, we use 0.627 ± 0.26, otherwise variations of one standard deviation are considered. These variations imply not only

modifications of the B^+ and B^0 relative fractions, but also of the κ-factor distributions and t-efficiency functions, which are considered in the estimation of the systematic uncertainty.

Fakes background in semileptonic samples: Two systematic sources are addressed regarding the description of the fakes background: its fraction and the shape of its proper time template. The latter is obtained from a fit to the fakes lepton sample, while the former is determined from a fit to the m_{Dl} distribution. Variations of both the fraction and the shape of this background component within the uncertainties obtained in these fits are propagated into contributions to the systematic uncertainty.

Proper time uncertainty calibration, S_t: The raw values of the proper decay time uncertainty σ_t are in general underestimated, and a scaling correction needs to be applied (5.10). Whereas the scale factors S_t for the $J/\psi K$ samples are determined as an integral part of the fit (5.33), for the $D\pi(\pi\pi)$ and DlX samples such direct determination is not attainable, because there is no prompt peak in proper time in these samples. For the t-biased hadronic and semileptonic samples we thus apply a fixed, common scale factor of 1.40 to the t resolution. This value is motivated by the $J/\psi K$ samples scale factor results of ~ 1.35, and by those obtained in Section 5.7. The fits are repeated by shifting this value, and the variations considered are $[1.0, 1.8]$ and $[1.3, 1.5]$ for the hadronic and semileptonic samples, respectively. The largest variations observed in the fit parameters (relative to the results with nominal fit conditions) are taken as the associated systematic uncertainty.

Construction of t-efficiency curves, $\mathcal{E}(t)$: The trigger selection and reconstruction requirements induce biasing effects in the proper decay time distribution which are accounted for by the t-efficiency function. This is constructed for each nominal signal mode from Monte Carlo samples which include trigger simulation and to which the sample selection is applied. These samples are modified at a time by different criteria. The B mesons lifetime input values are shifted by ± 1 standard deviation [1]. An observed difference between SVX and SVT d_0 residuals in Monte Carlo and data motivates the introduction of an extra smearing, of $\sim 12\mu$m, of the impact parameter d_0 before SVT-trigger confirmation. Tuning of resolution and efficiency for hits in the innermost silicon layer is applied to the simulation, which was demonstrated to further improve the agreement with data of L_{xy} and $\sigma_{L_{xy}}$ distributions. Corresponding t-efficiency curves are derived at a time and the combined fits repeated.

Discretized κ-factor distributions, $\mathcal{F}(\kappa)$: The proper time model for signal components in semileptonic samples involves the smearing with Monte Carlo-based κ-factor distributions, which are used to describe the effects of partial B reconstruction. About 500 variations of those histograms are produced by randomly modifying the entry of each bin according to its statistical uncertainty. The fits are repeated using such modified κ-factor distributions, and the associated systematic uncertainty contribution is taken as the width

of the distribution of fit values for each parameter.

Dilution templates, $L_{\mathcal{D}}$: The likelihood factors $L_{\mathcal{D}}$, which estimate for each OST algorithm the overall probability for a signal or background event to have a given predicted dilution, are realized by mass-sideband (subtracted) distributions. The bins of these template histograms are not always highly populated due to relatively low sample sizes and/or low tagging efficiencies. Sets of about 300 modified distributions each are generated, for every original template, by Gaussian-smearing the contents of each bin according to its statistical uncertainty. The associated systematic uncertainty is taken as the Gaussian-width of the distribution of fitted parameter results.

Additional effects evaluated for exclusive samples: The following sources are further evaluated for the exclusively reconstructed samples. In the high statistics exclusive $D\pi$ samples, the signal component is described in mass space by a double Gaussian, as it is observed to provide a better fit probability than the simpler single Gaussian model. As a systematic check, an alternative mass model involving a Breit-Wigner and Gaussian shape is used. For the Λ_b^0 component, appearing as part of the $B^0 \to D^-\pi^+$ sample, the SST dilution is fixed to 16% in the nominal fit. Although this quantity has not yet been measured, we have surmised that it is similar to the signal dilutions of B^+ and B^0. An associated systematic uncertainty is thus evaluated by repeating the primary fit with the Λ_b^0 SST dilution set to 12% and 20%. To cover the possibility that flavor tagging is slightly different among the modes, we compare the nominal fit with an alternative where all signal tagging efficiencies (SST, OSTs) and SST dilutions are fixed to the values found in individual fits of the modes. The background tagging efficiency parameters are fixed in the combined fit to the values found in individual fits for each of the analyzed decay modes. The statistical parameter uncertainties observed in these fits are used to generate sets of corrections to the nominal parameter values. The combined fit is repeated about 250 times, once for each such set, and the Gaussian-width of the distributions of fit results of the parameters of interest are taken as corresponding systematic uncertainty. In order to take into account correlation effects arising from the usage of the SST in the determination of the OST scale factors, the combined fit is repeated excluding SST information. We examine the shift in dilution scale factors between the two fits. Clearly, these variations are undefined for the SST dilutions, since they are not included in the varied fit. These are not applicable to Δm_d either, because its value is expected to change with the removal of SST information. However, for the dilution scale factor parameters, we opt to include the mentioned shifts as rather conservative estimates of possible correlation effects.

For some of the contributing systematic uncertainty sources above described, the estimation method adopted is based on an inherent statistical procedure. This may deliberately

result on a potential inflation of those particular systematic contributions. Regardless, we emphasize that the final measurement of the main parameters, summarized in Table 7.2, is dominated by statistical uncertainty.

The contributions to the systematic uncertainties are compiled in Tables 7.4 and 7.5, for the fully and the partially reconstructed combined modes. Among the prominent sources are the dilution templates, particularly for lower efficiency taggers and smaller size samples. This provides in effect the dominant contributions for the fully reconstructed samples. We mention in passing that the size of these contributions is, however, expected to decrease with the increased statistics of the data samples. For the partially reconstructed modes, the dominant source of systematic uncertainties for the dilution scale factors is the fraction of the fakes background, especially for the high statistics, low purity taggers. For the Δm_d parameter, the predominant systematic uncertainty comes from the Dl signal composition. Although the present Δm_d measurement is dominated by statistical uncertainty, sample composition effects are ineluctably expected to limit the precision of future measurements in semileptonic decays.

source	relative uncertainty [%]							
	S_{SMT}	S_{SET}	S_{JVX}	S_{JJP}	S_{JPT}	$\mathcal{D}^{B^+}_{SST}$	$\mathcal{D}^{B^0}_{SST}$	Δm_d
combinatorial bckg.	0.3	0.6	0.7	0.7	0.7	0.3	1.6	0.6
physics bckg. level								
K^{*0} swap in $J/\psi K^{*0}$	< 0.1	< 0.1	0.3	0.4	0.2	< 0.1	0.1	< 0.1
Λ^0_b and B_s in $D^-\pi^+$	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.4	0.1
Cabibbo-suppressed	0.4	0.3	0.8	0.1	0.6	< 0.1	0.6	0.2
signal shape for $D\pi$	< 0.1	0.4	0.5	< 0.1	< 0.1	< 0.1	0.2	0.2
scale factor on σ_t	< 0.1	< 0.1	0.2	< 0.1	< 0.1	< 0.1	0.2	0.2
t-efficiency function	< 0.1	< 0.1	0.6	0.3	1.0	< 0.1	0.4	0.1
Dilution templates	2.2	3.9	3.3	1.0	2.5	0.5	1.5	0.7
Tagging efficiencies	0.5	0.4	0.3	< 0.1	0.4	0.1	0.6	0.5
Λ^0_b SST dilution in $D\pi$	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	—	< 0.1	< 0.1
SST removal	0.6	1.4	1.1	2.6	7.2	—	—	—
total	2.4	4.2	3.7	2.9	7.8	0.6	2.4	1.1

Table 7.4: Summary of systematic uncertainties in fully reconstructed modes.

source	relative uncertainty [%]					
	S_{SMT}	S_{SET}	S_{JVX}	S_{JJP}	S_{JPT}	Δm_d
combinatorial background	0.1	0.1	0.1	0.2	0.3	0.2
physics background	0.1	0.1	0.1	< 0.1	0.1	0.8
fakes background fraction	0.1	0.6	1.2	2.7	5.0	0.1
fakes background shape	0.1	0.1	0.3	0.3	0.4	0.1
signal composition	0.3	0.2	0.3	0.5	0.1	2.6
scale factor on σ_t	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1
t-efficiency function	0.2	0.1	0.2	0.2	0.4	0.2
κ-factor binning	0.2	0.1	0.1	0.1	0.3	0.1
dilution templates	1.3	1.4	0.7	0.5	0.4	0.5
total	1.4	1.6	1.4	2.8	5.1	2.8

Table 7.5: Summary of systematic uncertainties in partially reconstructed modes.

7.7 Résumé

The likelihood description of the data samples has been extended to include b-flavor information, and has been applied to the samples of B^0 and B^+ meson candidates in flavor oscillation and tagging measurements.

The opposite-side and same-side tagging methods – abbreviated OST and SST – are described in Chapter 6. We apply these techniques to compare the flavor, *i.e.* B or $\bar{B}$ state, of our $B^{+,0}$ candidates at production and decay times. The OST is used for tagging the semileptonic samples, while both OST and SST are applied to the fully reconstructed modes. The OST algorithms are combined in an exclusive fashion, such that for a given candidate only the decision provided by one selected algorithm is used. The combination of OST and SST information is derived, and also implemented in the fitting model.

The samples description in mass and proper decay time spaces achieved in Chapter 5 is augmented to incorporate the candidates flavor tagging information. For neutral B candidates, the corresponding likelihood signal component contains an additional time-dependent oscillating term which describes mixing. The B^0 oscillation frequency is determined directly as a fit parameter. The combined fit result is

$$\Delta m_d = 0.522 \pm 0.017 \text{ ps}^{-1} .$$

The measurement is in good agreement with the world average value, and the uncertainty is still dominated by the statistical contributions.

Predicted OST dilution values are assigned to the individual candidates based on properties of the event, and an overall dilution calibration factor is introduced as a fit parameter for each algorithm. For the SST, the average dilutions are measured directly as parameters of the fit. The average effectiveness of the two classes of tagging methods in the combined samples is:

method	tagging power, $\epsilon\mathcal{D}^2$ [%]
opposite-side	1.55 ± 0.08
same-side, B^+	2.77 ± 0.30
same-side, B^0	1.07 ± 0.30

The distinct behavior of the SST for different B meson species is emphasized. The combined OST performance is that obtained with the exclusive algorithm combination employed.

The treatment of flavor tagging and mixing here described in the context of the $B^{+,0}$ samples serves as the basis for the study of flavor oscillations in the B_s system, to be carried out in the following chapters. The achieved taggers' calibration also constitutes an important factor for setting reliable exclusion conditions on the B_s oscillation frequency.